



Das Arbeiten mit Künstlicher Intelligenz (KI) wird in immer mehr Unternehmen und Institutionen zum Standard. Doch wirft dies zahlreiche ethische Probleme auf – und ist sogar mit ganz konkreten Gefahren verbunden.

Foto: picture alliance/dpa/Atos Information Technology GmbH

„Die Debatte wird falsch geführt“

Der Bayreuther KI-Professor Niklas Kühl äußert sich zu existenziellen Risiken der Künstlichen Intelligenz. Er besitzt eine klare Meinung, was (Welt-)Untergangsszenarien durch KI betrifft: Dies sei Spekulation und keine Wissenschaft.

BAYREUTH. Anthropic, eine der bedeutendsten Firmen zur Entwicklung Künstlicher Intelligenz, warnt vor existenziellen Risiken für die Menschheit. Gerät die Künstliche Intelligenz außer Kontrolle? Wie kann man die Risiken minimieren? Wir fragten nach bei Niklas Kühl, Bayreuther Professor für humanzentrische KI.

Anthropic warnt vor existenziellen Risiken für die Menschheit durch KI. Übertrieben?

Prof. Niklas Kühl: Es gibt reale Risiken, und einige davon könnten im Extremfall existenziell sein. Die Debatte wird aber falsch geführt: Über die Probleme, die KI schon heute verursacht, wird viel zu wenig gesprochen. Stattdessen wird sofort das sehr unwahrscheinliche Weltuntergangsszenario bemittelt.

Insider behaupten, die Künstliche Intelligenz könne die Menschheit binnen eines Jahrzehnts auslöschen. Nur Panikmache?



N. Kühl Foto: privat

Solche Aussagen beruhen meist auf „P(doom)“-Schätzungen, also gefühlten Wahrscheinlichkeiten für den Weltuntergang. Da liegen Fachleute weit auseinander. Ganz ausschließen kann man nichts, aber das sind Gedankenexperimente ohne belastbare Grundlage. Für mich ist das Spekulation, keine Wissenschaft.

Was könnte denn schlimmstenfalls passieren?

Ein Zusammenbruch aller digital gestützten Dienste. Und das ist heute praktisch alles: Zahlungsverkehr, Lieferketten, Kommunikation. Ein weiteres, wenn auch unwahrscheinliches Risiko ist der Einsatz von KI bei der Entwicklung oder Steuerung von Waffen.

Wie hoch ist die Wahrscheinlichkeit für Worst Case?

Sehr gering, seriös beziffern kann sie niemand. Mein Punkt ist ein anderer: Aufmerksamkeit und politische Energie sind begrenzt. Wenn die Debatte von Untergangsszenarien bestimmt wird, fehlt diese Energie dort, wo wir tatsächlich etwas bewirken können: bei einer wirksamen KI-Regulierung, bei einer stärkeren Rolle Europas in der KI und digitalen Souveränität und bei drängenden Problemen wie Klimawandel, wachsender Ungleichheit und demokratiefeindlichen Tendenzen.

Die KI-Entwicklung geht stürmisch voran. Gewinnstreben vor Risikobewusstsein?

Ja, aber das ist nichts Neues, das kennen wir aus jeder Technologiewelle. Und man sollte genau hinschauen: Wenn ein Unternehmen vor dem Börsengang existenzielle Risiken in seinen Prospekt schreibt, ist das auch juristi-

sche Absicherung gegenüber Anlegern und nicht nur Verantwortungsbewusstsein. Oder sogar Werbung: Schaut her, wie gefährlich wir sind!

Wird zu leichtfertig programmiert?

Vor allem wird so getan, als seien die Firmen nicht schuld, wenn etwas schiefgeht. Überall sonst wird man zur Rechenschaft gezogen, warum hier nicht? Es fehlt an echter Rechenschaftspflicht.

Der Gesetzgeber müsste bremsen? Kann er überhaupt?

Er kann, und er tut es auch: Die EU und China haben klare Regeln. Aber KI ist global, also braucht es auch globale Regulierung. Ein einfacher erster Schritt wäre, die Anbieter endlich für die Schäden haftbar zu machen, die ihre Systeme anrichten.

„Über die Probleme, die KI schon heute verursacht, wird viel zu wenig gesprochen.“

Niklas Kühl KI-Experte

US-Präsident Donald Trump hat wenig Interesse, die KI-Entwicklung stärker staatlich zu kontrollieren?

KI ist die einzige „winning story“, die den USA derzeit noch bleibt. Regulierung passt nicht in diese Erzählung, und eine differenzierte und seriöse Auseinandersetzung mit der Technologie ist von dieser Regierung nicht zu erwarten.

Anthropic steht vor dem Börsengang. Mit über zwei Billionen Dollar wird die

Börse das KI-Labor bewerten, wird in Aussicht gestellt. Das viele Geld wird die KI-Entwicklung noch mehr anheizen?

Warten wir den Börsengang erst einmal ab. Laut dem Börsenprospekt, aus dem Reuters berichtet, hat Anthropic 2025 rund 4,6 Milliarden Dollar Umsatz gemacht, aber einen operativen Verlust von gut acht Milliarden. Allein für Rechenleistung und Infrastruktur flossen 7,3 Milliarden, also etwa 1,60 Dollar pro Dollar Umsatz. Dazu kommen Zusagen über 518 Milliarden Dollar für Rechenkapazität, von denen rund 80 Prozent nicht kündbar sind. Der Umsatz wächst zwar rasant, zuletzt rund 11,5 Milliarden in einem einzigen Quartal. Eine Bewertung von zwei Billionen ist trotzdem eine gewaltige Wette darauf, dass dieses Wachstum über Jahre anhält und die enormen Kosten irgendwann trägt.

Ihre Prognose: Die KI - Fluch oder Segen?

Unreguliert und ohne Rechenschaftspflicht: ein Fluch. Mit klaren Regeln und echter Haftung kann KI ein Segen sein.

Das Gespräch führte Roland Töpfer

Zur Person

Niklas Kühl ist Professor und Inhaber des Lehrstuhls für Wirtschaftsinformatik und humanzentrische Künstliche Intelligenz an der Universität Bayreuth. Kühl forscht an der Schnittstelle von Künstlicher Intelligenz und deren verantwortungsvoller Anwendung in Wirtschaft und Gesellschaft. Im Zentrum steht dabei die Zusammenarbeit zwischen Menschen und KI.